GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:26:14 ; Search time 53 Seconds
(without alignments)
95.960 Million cell updates/sec

Title: US-09-747-029B-17
Perfect score: 105
Sequence: 1 QDTIVGWGCDSXGCRPGQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 481716

Minimum DB seq length: 0
Maximum DB seq length: 18

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	103	98.1	18	4	AAE07230	Aae07230 IGP1685 p
2	83	79.0	14	4	AAE07231	Aae07231 IGP1686 p
3	60	57.1	18	4	AAE07225	Aae07225 IGP1650 p
4	47	44.8	16	4	AAB76391	Aab76391 ErbB2 bin
5	47	44.8	16	4	AAB76385	Aab76385 ErbB2 bin
6	44	41.9	14	4	AAE07229	Aae07229 IGP1684 p
7	44	41.9	16	4	AAB76382	Aab76382 ErbB2 bin
8	43	41.0	14	4	AAE07227	Aae07227 IGP1676 p
9	43	41.0	18	4	AAE07223	Aae07223 IGP1648 p
10	42	40.0	16	4	AAB76392	Aab76392 ErbB2 bin
11	42	40.0	16	4	AAB76387	Aab76387 ErbB2 bin
12	42	40.0	17	4	AAB76369	Aab76369 ErbB2 bin
13	42	40.0	17	4	AAB76370	Aab76370 ErbB2 bin
14	42	40.0	18	4	AAE07221	Aae07221 IGP1646 p
15	41	39.0	16	4	AAB76383	Aab76383 ErbB2 bin
16	41	39.0	16	4	AAB76390	Aab76390 ErbB2 bin
17	41	39.0	17	4	AAB76354	Aab76354 ErbB2 bin
18	40	38.1	14	4	AAE07228	Aae07228 IGP1687 p
19	39	37.1	11	2	AAW09559	Aaw09559 Thrombopo
20	39	37.1	11	2	AAW36710	Aaw36710 Thrombopo
21	39	37.1	11	4	AAU25929	Aau25929 Human thr
22	39	37.1	18	4	AAE07224	Aae07224 IGP1649 p
23	38	36.2	14	2	AAY08356	Aay08356 Cysteine
24	38	36.2	15	4	AAJ03265	Aaj03265 Hepatitis
25	38	36.2	15	4	AAJ03570	Aaj03570 Hepatitis
26	38	36.2	15	4	AAJ03655	Aaj03655 Hepatitis
27	38	36.2	17	4	AAB76353	Aab76353 ErbB2 bin
28	38	36.2	17	4	AAB76373	Aab76373 ErbB2 bin
29	38	36.2	18	4	AAE07220	Aae07220 IGP1611 p
30	37	35.2	12	4	AAE06030	Aae06030 Dodecamer
31	37	35.2	15	2	AAR93669	Aar93669 HIV princ
32	37	35.2	16	4	AAB76384	Aab76384 ErbB2 bin
33	37	35.2	17	2	AAR06085	Aar06085 Immunorea
34	36	34.3	10	5	ABG98831	Abg98831 F protein
35	36	34.3	10	5	ABG98832	Abg98832 F protein
36	36	34.3	10	5	ABG98833	Abg98833 F protein
37	36	34.3	16	4	AAB76388	Aab76388 ErbB2 bin
38	36	34.3	16	4	AAB76386	Aab76386 ErbB2 bin
39	36	34.3	16	4	AAB76381	Aab76381 ErbB2 bin
40	36	34.3	16	4	AAG73166	Aag73166 Protease
41	36	34.3	17	4	AAM15636	Aam15636 Peptide #
42	36	34.3	17	4	AAB76372	Aab76372 ErbB2 bin
43	36	34.3	17	4	AAB76368	Aab76368 ErbB2 bin
44	36	34.3	17	4	AAB76371	Aab76371 ErbB2 bin
45	36	34.3	17	4	AAB76375	Aab76375 ErbB2 bin

ALIGNMENTS

RESULT 1
AAE07230
ID AAE07230 standard; peptide; 18 AA.
XX
AC AAE07230;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1685 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1685; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .18
FT /note= "Biotinylated residues"
FT Disulfide-bond 9. .14
FT Modified-site 12
FT /note= "Citrulline"
XX
PN WO200146222-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-EP013037.
XX
PR 21-DEC-1999; 99EP-00870280.
PR 08-SEP-2000; 2000EP-00870195.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Union A, Moereels H, Meheus L;
XX
DR WPI; 2001-496657/54.
XX
PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT comprises citrulline residue between 2 cysteine residues and is
PT specifically recognized by autoimmune antibodies from patients suffering
PT from rheumatoid arthritis.
XX
PS Claim 9; Page 42; 53pp; English.
XX
CC The present sequence is a cyclic synthetic biotinylated peptide, IGP1685.
CC The peptide comprises a citrulline residue between 2 cysteine residues
CC and is specifically recognised by autoimmune antibodies from patients

CC suffering from rheumatoid arthritis. The peptide comprises amino acids
CC involved in side chain interactions which is essential for the formation
CC of three-dimensional structure of the peptide. The peptide of the
CC invention is useful as a medicament to treat autoimmune diseases,
CC preferably rheumatoid arthritis. It is also useful for treating
CC autoimmune diseases by increasing the size of antigen-immune complexes to
CC improve clearance of the formed immune complexes and for the preparation
CC of a medicament for oral or nasal administration to treat autoimmune
CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC to the peptide
XX
SQ Sequence 18 AA;

Query Match 98.1%; Score 103; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDTIVGWGCDSXGCRPGQ 18
||||||||||||||||||
Db 1 QDTIVGWGCDSXGCRPGQ 18

RESULT 2
AAE07231
ID AAE07231 standard; peptide; 14 AA.
XX
AC AAE07231;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1686 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1686; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .14
FT /note= "Biotinylated residues"
FT Disulfide-bond 9. .14
FT Modified-site 12
FT /note= "Citrulline"
XX
PN WO200146222-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-EP013037.
XX
PR 21-DEC-1999; 99EP-00870280.
PR 08-SEP-2000; 2000EP-00870195.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Union A, Moereels H, Meheus L;
XX
DR WPI; 2001-496657/54.
XX
PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT comprises citrulline residue between 2 cysteine residues and is
PT specifically recognized by autoimmune antibodies from patients suffering
PT from rheumatoid arthritis.
XX
PS Claim 9; Page 42; 53pp; English.
XX
CC The present sequence is a cyclic synthetic biotinylated peptide, IGP1686.
CC The peptide comprises a citrulline residue between 2 cysteine residues
CC and is specifically recognised by autoimmune antibodies from patients
CC suffering from rheumatoid arthritis. The peptide comprises amino acids
CC involved in side chain interactions which is essential for the formation
CC of three-dimensional structure of the peptide. The peptide of the
CC invention is useful as a medicament to treat autoimmune diseases,
CC preferably rheumatoid arthritis. It is also useful for treating
CC autoimmune diseases by increasing the size of antigen-immune complexes to
CC improve clearance of the formed immune complexes and for the preparation
CC of a medicament for oral or nasal administration to treat autoimmune
CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC to the peptide
XX
SQ Sequence 14 AA;

Query Match 79.0%; Score 83; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VGWGCDSXGCRPGQ 18
||||||||||||||
Db 1 VGWGCDSXGCRPGQ 14

RESULT 3
AAE07225
ID AAE07225 standard; peptide; 18 AA.
XX
AC AAE07225;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1650 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1650; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .18
FT /note= "Biotinylated residues"
FT Disulfide-bond 9. .14
FT Modified-site 11
FT /note= "Citrulline"
FT Modified-site 12
FT /note= "Citrulline"
XX
PN WO200146222-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-EP013037.
XX
PR 21-DEC-1999; 99EP-00870280.
PR 08-SEP-2000; 2000EP-00870195.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Union A, Moereels H, Meheus L;
XX
DR WPI; 2001-496657/54.
XX
PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT comprises citrulline residue between 2 cysteine residues and is
PT specifically recognized by autoimmune antibodies from patients suffering
PT from rheumatoid arthritis.
XX
PS Claim 9; Page 42; 53pp; English.
XX
CC The present sequence is a cyclic synthetic biotinylated peptide, IGP1650.
CC The peptide comprises a citrulline residue between 2 cysteine residues
CC and is specifically recognised by autoimmune antibodies from patients
CC suffering from rheumatoid arthritis. The peptide comprises amino acids
CC involved in side chain interactions which is essential for the formation
CC of three-dimensional structure of the peptide. The peptide of the
CC invention is useful as a medicament to treat autoimmune diseases,

CC preferably rheumatoid arthritis. It is also useful for treating
CC autoimmune diseases by increasing the size of antigen-immune complexes to
CC improve clearance of the formed immune complexes and for the preparation
CC of a medicament for oral or nasal administration to treat autoimmune
CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC to the peptide
XX
SQ Sequence 18 AA;

Query Match 57.1%; Score 60; DB 4; Length 18;
Best Local Similarity 70.6%; Pred. No. 0.32;
Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QDTIVGWGCDSXGCRPG 17
||||||||||||||||||
Db 1 QDTIHGHPCSXXGCRPG 17

RESULT 4
AAB76391
ID AAB76391 standard; peptide; 16 AA.
XX
AC AAB76391;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 42.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 16 AA;

Query Match 44.8%; Score 47; DB 4; Length 16;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGCR 15
|||| |||
Db 5 GWGCIGPGCR 14

RESULT 5
AAB76385
ID AAB76385 standard; peptide; 16 AA.
XX
AC AAB76385;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 36.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 16 AA;

Query Match 44.8%; Score 47; DB 4; Length 16;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGCR 15
|||| |||
Db 5 GWGCIGPGCR 14

RESULT 6
AAE07229
ID AAE07229 standard; peptide; 14 AA.
XX
AC AAE07229;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1684 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1684; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .14
FT /note= "Biotinylated residues"
FT Disulfide-bond 9. .16
FT Modified-site 12
FT /note= "Citrulline"
XX
PN WO200146222-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-EP013037.
XX
PR 21-DEC-1999; 99EP-00870280.
PR 08-SEP-2000; 2000EP-00870195.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Union A, Moereels H, Meheus L;
XX
DR WPI; 2001-496657/54.
XX
PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT comprises citrulline residue between 2 cysteine residues and is
PT specifically recognized by autoimmune antibodies from patients suffering
PT from rheumatoid arthritis.
XX
PS Claim 9; Page 42; 53pp; English.
XX
CC The present sequence is a cyclic synthetic biotinylated peptide, IGP1684.
CC The peptide comprises a citrulline residue between 2 cysteine residues
CC and is specifically recognised by autoimmune antibodies from patients
CC suffering from rheumatoid arthritis. The peptide comprises amino acids
CC involved in side chain interactions which is essential for the formation
CC of three-dimensional structure of the peptide. The peptide of the
CC invention is useful as a medicament to treat autoimmune diseases,
CC preferably rheumatoid arthritis. It is also useful for treating
CC autoimmune diseases by increasing the size of antigen-immune complexes to
CC improve clearance of the formed immune complexes and for the preparation
CC of a medicament for oral or nasal administration to treat autoimmune
CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC to the peptide
XX
SQ Sequence 14 AA;

Query Match 41.9%; Score 44; DB 4; Length 14;
Best Local Similarity 76.9%; Pred. No. 34;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGCRPGQ 18
|| |||||| | ||
Db 2 GHGCDSXGHRCGQ 14

RESULT 7
AAB76382
ID AAB76382 standard; peptide; 16 AA.
XX
AC AAB76382;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 33.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 16 AA;

Query Match 41.9%; Score 44; DB 4; Length 16;
Best Local Similarity 60.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGCR 15
|||| ||:
Db 5 GWGCIGPGCK 14

RESULT 8
AAE07227
ID AAE07227 standard; peptide; 14 AA.
XX
AC AAE07227;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1676 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1676; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;

```
KW   systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .14
FT                   /note= "Biotinylated residues"
FT   Disulfide-bond  9. .14
FT   Modified-site   11
FT                   /note= "Citrulline"
FT   Modified-site   12
FT                   /note= "Citrulline"
XX
PN   WO200146222-A2.
XX
PD   28-JUN-2001.
XX
PF   20-DEC-2000; 2000WO-EP013037.
XX
PR   21-DEC-1999;   99EP-00870280.
PR   08-SEP-2000; 2000EP-00870195.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Union A,  Moereels H,  Meheus L;
XX
DR   WPI; 2001-496657/54.
XX
PT   New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT   comprises citrulline residue between 2 cysteine residues and is
PT   specifically recognized by autoimmune antibodies from patients suffering
PT   from rheumatoid arthritis.
XX
PS   Claim 9; Page 42; 53pp; English.
XX
CC   The present sequence is a cyclic synthetic biotinylated peptide, IGP1676.
CC   The peptide comprises a citrulline residue between 2 cysteine residues
CC   and is specifically recognised by autoimmune antibodies from patients
CC   suffering from rheumatoid arthritis. The peptide comprises amino acids
CC   involved in side chain interactions which is essential for the formation
CC   of three-dimensional structure of the peptide. The peptide of the
CC   invention is useful as a medicament to treat autoimmune diseases,
CC   preferably rheumatoid arthritis. It is also useful for treating
CC   autoimmune diseases by increasing the size of antigen-immune complexes to
CC   improve clearance of the formed immune complexes and for the preparation
CC   of a medicament for oral or nasal administration to treat autoimmune
CC   diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC   to the peptide
XX
SQ   Sequence 14 AA;

  Query Match             41.0%;  Score 43;  DB 4;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 46;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 GWGCDSXGCRPG 17
              |  |  ||||||
Db          2 GHPCSXXGCRPG 13


RESULT 9
AAE07223
ID   AAE07223 standard; peptide; 18 AA.
XX
AC   AAE07223;
XX
DT   06-NOV-2001  (first entry)
XX
DE   IGP1648 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW   Synthetic peptide; cyclic; IGP1648; autoimmune antibody;
KW   rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW   systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .18
FT                   /note= "Biotinylated residues"
FT   Disulfide-bond  9. .16
FT   Modified-site   11
FT                   /note= "Citrulline"
FT   Modified-site   12
FT                   /note= "Citrulline"
XX
PN   WO200146222-A2.
XX
PD   28-JUN-2001.
XX
PF   20-DEC-2000; 2000WO-EP013037.
XX
PR   21-DEC-1999;   99EP-00870280.
PR   08-SEP-2000; 2000EP-00870195.
XX
PA   (INNO-) INNOGENETICS NV.
XX
PI   Union A,  Moereels H,  Meheus L;
XX
DR   WPI; 2001-496657/54.
XX
PT   New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT   comprises citrulline residue between 2 cysteine residues and is
PT   specifically recognized by autoimmune antibodies from patients suffering
PT   from rheumatoid arthritis.
XX
PS   Claim 9; Page 42; 53pp; English.
XX
CC   The present sequence is a cyclic synthetic biotinylated peptide, IGP1648.
CC   The peptide comprises a citrulline residue between 2 cysteine residues
CC   and is specifically recognised by autoimmune antibodies from patients
CC   suffering from rheumatoid arthritis. The peptide comprises amino acids
CC   involved in side chain interactions which is essential for the formation
CC   of three-dimensional structure of the peptide. The peptide of the
CC   invention is useful as a medicament to treat autoimmune diseases,
CC   preferably rheumatoid arthritis. It is also useful for treating
CC   autoimmune diseases by increasing the size of antigen-immune complexes to
CC   improve clearance of the formed immune complexes and for the preparation
CC   of a medicament for oral or nasal administration to treat autoimmune
CC   diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC   to the peptide
XX
SQ   Sequence 18 AA;

  Query Match             41.0%;  Score 43;  DB 4;  Length 18;
  Best Local Similarity   61.1%;  Pred. No. 58;
  Matches   11;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 QDTIVGWGCDSXGCRPGQ 18
              ||||  |  |  || | ||
Db          1 QDTIHGHPCSXXGHRCGQ 18


RESULT 10
AAB76392
ID   AAB76392 standard; peptide; 16 AA.
XX
AC   AAB76392;
XX
DT   10-APR-2001  (first entry)
XX
DE   ErbB2 binding peptide amino acid sequence SEQ ID 43.
XX
KW   Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW   metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW   Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW   Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
```

XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 16 AA;

Query Match 40.0%; Score 42; DB 4; Length 16;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGC 14
|||| ||
Db 5 GWGCIGPGC 13

RESULT 11
AAB76387
ID AAB76387 standard; peptide; 16 AA.
XX
AC AAB76387;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 38.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 16 AA;

Query Match 40.0%; Score 42; DB 4; Length 16;
Best Local Similarity 66.7%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 6 GWGCDSXGC 14
|||| ||
Db 5 GWGCIGPGC 13

RESULT 12
AAB76369
ID AAB76369 standard; peptide; 17 AA.
XX
AC AAB76369;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 20.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human

PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 17 AA;

Query Match 40.0%; Score 42; DB 4; Length 17;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          6 GWGCDSXGC 14
              ||||   ||
Db          2 GWGCIGPGC 10
```

RESULT 13
AAB76370
ID AAB76370 standard; peptide; 17 AA.
XX
AC AAB76370;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 21.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX
SQ Sequence 17 AA;

Query Match 40.0%; Score 42; DB 4; Length 17;
Best Local Similarity 66.7%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          6 GWGCDSXGC 14
              ||||   ||
Db          2 GWGCIGPGC 10
```

RESULT 14
AAE07221
ID AAE07221 standard; peptide; 18 AA.
XX
AC AAE07221;
XX
DT 06-NOV-2001 (first entry)
XX
DE IGP1646 peptide for diagnosis and treatment of rheumatoid arthritis.
XX
KW Synthetic peptide; cyclic; IGP1646; autoimmune antibody;
KW rheumatoid arthritis; therapy; autoimmune disease; antirheumatic;
KW systemic hyporesponsiveness; immunosuppressive; antiarthritic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1. .18
FT /note= "Biotinylated residues"
FT Disulfide-bond 9. .16
FT Modified-site 12
FT /note= "Citrulline"
XX
PN WO200146222-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-EP013037.
XX
PR 21-DEC-1999; 99EP-00870280.
PR 08-SEP-2000; 2000EP-00870195.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Union A, Moereels H, Meheus L;
XX
DR WPI; 2001-496657/54.
XX
PT New peptides, useful for diagnosing and treating rheumatoid arthritis,
PT comprises citrulline residue between 2 cysteine residues and is
PT specifically recognized by autoimmune antibodies from patients suffering
PT from rheumatoid arthritis.
XX
PS Claim 9; Page 42; 53pp; English.
XX
CC The present sequence is a cyclic synthetic biotinylated peptide, IGP1646.
CC The peptide comprises a citrulline residue between 2 cysteine residues
CC and is specifically recognised by autoimmune antibodies from patients
CC suffering from rheumatoid arthritis. The peptide comprises amino acids
CC involved in side chain interactions which is essential for the formation

CC of three-dimensional structure of the peptide. The peptide of the
CC invention is useful as a medicament to treat autoimmune diseases,
CC preferably rheumatoid arthritis. It is also useful for treating
CC autoimmune diseases by increasing the size of antigen-immune complexes to
CC improve clearance of the formed immune complexes and for the preparation
CC of a medicament for oral or nasal administration to treat autoimmune
CC diseases by inducing a state of systemic hyporesponsiveness or tolerance
CC to the peptide
XX
SQ Sequence 18 AA;

Query Match 40.0%; Score 42; DB 4; Length 18;
Best Local Similarity 64.7%; Pred. No. 78;
Matches 11; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 QDTIVGWGCDSXGCRPG 17
              |||| |  | ||| | |
Db          1 QDTIHGHPCSSXGHRCG 17
```

RESULT 15
AAB76383
ID AAB76383 standard; peptide; 16 AA.
XX
AC AAB76383;
XX
DT 10-APR-2001 (first entry)
XX
DE ErbB2 binding peptide amino acid sequence SEQ ID 34.
XX
KW Human; erbB2; HER2; cancer; nervous system disease; stroke; ischaemia;
KW metabolic disorder; nutritional deficiency; Alzheimer's disease;
KW Parkinson's disease; epilepsy; multiple sclerosis; Huntington's chorea;
KW Down's syndrome; nerve deafness; Meniere's disease; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200101748-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US018283.
XX
PR 02-JUL-1999; 99US-0142232P.
XX
PA (GETH) GENENTECH INC.
XX
PI Dennis MS;
XX
DR WPI; 2001-123048/13.
XX
PT Non-naturally occurring peptide ligands which compete for binding human
PT erB2 gene products, useful for treating e.g. Alzheimer's disease,
PT multiple sclerosis and diabetic neuropathy.
XX
PS Disclosure; Fig 16; 116pp; English.
XX
CC This invention relates to non-naturally occurring peptide ligands which
CC bind to the human erbB2 gene product ErbB2 (also known as HER2). Peptides
CC represented in AAB76350 - AAB76420 and AAB76432 - AAB76509 are examples
CC of the ErbB2 binding ligands of the invention. Sequences AAB76421 -
CC AAB76431 represent antibody Fc amino acid sequences used in the isolation
CC of the peptides of the invention. The peptides compete for binding ErbB2
CC with naturally occurring ligands, and may be used to treat disorders
CC characterized by over expression of HER2/ErbB2 such as cancers, diseases
CC of the nervous system, musculature and epithelia, e.g. nervous system
CC damage resulting from trauma, surgery, strokes, ischaemia, infection,
CC metabolic disorders, nutritional deficiency or toxic agents. In
CC particular the synthetic peptide ligands may be used to treat Alzheimer's
CC disease, Parkinson's disease, epilepsy, multiple sclerosis, Huntington's
CC chorea, Down's syndrome, nerve deafness, Meniere's disease and neuropathy
CC associated with diabetes
XX

SQ Sequence 16 AA;

Query Match 39.0%; Score 41; DB 4; Length 16;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 WGCDSXGCR 15
              |||   |||
Db          6 WGCIGPGCR 14
```

Search completed: March 13, 2004, 07:28:55
Job time : 54 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:26:14 ; Search time 20 Seconds
(without alignments)
86.572 Million cell updates/sec

Title: US-09-747-029B-17
Perfect score: 105
Sequence: 1 QDTIVGWGCDSXGCRPGQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 3214

Minimum DB seq length: 0
Maximum DB seq length: 18

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	27.6	15	2	D48394	major fat-globule
2	27	25.7	17	2	A61211	anantin - Streptom
3	26	24.8	10	1	RHPGG	gonadoliberin - pi
4	26	24.8	10	1	RHSHG	gonadoliberin - sh
5	25	23.8	15	2	I78838	flt3 ligand isofor
6	24	22.9	9	1	AKLQIM	locustamyoinhibiti
7	24	22.9	11	2	D45900	complement C3b rec
8	24	22.9	15	2	S39012	proteinase - Therm
9	24	22.9	17	2	A44560	terephthalate 1,2-
10	23	21.9	10	2	A60421	hypertrehalosemic
11	23	21.9	10	2	S08997	hypertrehalosemic
12	23	21.9	10	2	S08998	hypertrehalosemic
13	23	21.9	10	2	A26381	hypertrehalosemic
14	23	21.9	10	2	A31571	hypertrehalosemic/
15	23	21.9	10	2	B33995	hypotrehalosemic h
16	23	21.9	11	2	S33300	probable substance
17	23	21.9	14	2	PH0755	T-cell receptor be
18	23	21.9	16	2	I78533	gene agoute protei
19	22	21.0	8	2	I57018	gene Cftr protein
20	22	21.0	9	2	S39437	D-amino-acid oxida
21	22	21.0	10	1	RHAQ1	gonadoliberin I -
22	22	21.0	10	2	A13687	caerulein-like pep
23	22	21.0	11	2	S60354	retinal oxidase -
24	22	21.0	12	2	A49033	T-cell receptor de
25	22	21.0	13	2	PT0305	Ig heavy chain CRD
26	22	21.0	15	2	B56046	urinary tract ston
27	22	21.0	16	1	A49761	locustapyrokinin -
28	22	21.0	17	2	A61334	trypsin (EC 3.4.21
29	21.5	20.5	16	4	I79565	hypothetical TCL3/
30	21	20.0	9	2	A57444	neuropeptide Grb-A
31	21	20.0	11	2	B49164	chromogranin-B - r
32	21	20.0	12	2	I40663	bma protein - Clos
33	21	20.0	14	2	B61309	lutropin beta chai
34	21	20.0	15	2	S08301	epidermal growth f
35	21	20.0	16	2	D36912	hypothetical prote
36	21	20.0	17	2	PH1357	Ig heavy chain DJ
37	21	20.0	18	2	S21669	1H-4-oxoquinoline
38	21	20.0	18	2	B49048	T-cell receptor be
39	20	19.0	8	2	A31570	angiotensin-conver
40	20	19.0	8	2	JS0316	leucokinin VI - Ma
41	20	19.0	10	2	A61337	caerulein - frog (
42	20	19.0	11	2	PT0273	Ig heavy chain CRD
43	20	19.0	11	2	PH0940	T-cell receptor be
44	20	19.0	12	2	B49033	T-cell receptor de
45	20	19.0	13	1	XAVI9B	angiotensin-conver

ALIGNMENTS

RESULT 1
D48394
major fat-globule membrane protein GP 55 - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997
C;Accession: D48394
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pi
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: D48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131448)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homol

Query Match 27.6%; Score 29; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 5e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          8 GCDSXGC 14
              ||:  ||
Db          5 GCELNGC 11
```

RESULT 2
A61211
anantin - Streptomyces coerulescens
C;Species: Streptomyces coerulescens
C;Date: 03-May-1994 #sequence_revision 05-Apr-1995 #text_change 07-May-1999
C;Accession: A61211
R;Wyss, D.F.; Lahm, H.W.; Manneberg, M.; Labhardt, A.M.
J. Antibiot. 44, 172-180, 1991
A;Title: Anantin -- a peptide antagonist of the atrial natriuretic factor (ANF). II. De
A;Reference number: A61211; MUID:91185186; PMID:1826288
A;Accession: A61211
A;Molecule type: protein
A;Residues: 1-17 <WYS>
A;Note: the isopeptide linked residue 8 is shown as Asn rather than Asp
F;1-8/Cross-link: isopeptide amino end (Gly-Asn) #status experimental

Query Match 25.7%; Score 27; DB 2; Length 17;
Best Local Similarity 44.4%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 VGWGCDSXG 13
              :|||  :  |
Db          3 IGWGNNIFG 11
```

```
RESULT 3
RHPGG
gonadoliberin - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C;Accession: A01411
R;Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A;Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
A;Reference number: A90172; MUID:72114303; PMID:4946067
A;Accession: A01411
A;Molecule type: protein
A;Residues: 1-10 <BAB>
R;Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A;Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase metho
A;Reference number: A90176; MUID:72065376; PMID:4942726
A;Contents: annotation; synthesis
A;Note: the synthetic and natural hormones have the same physicochemical and biological
R;Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A;Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A;Reference number: A90175; MUID:72117544; PMID:4946275
A;Contents: annotation
A;Note: Trp-3 appears to be essential for biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             24.8%;  Score 26;  DB 1;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 9.7e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 SXGCRPG 17
               | | |||
Db          4 SYGLRPG 10


RESULT 4
RHSHG
gonadoliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C;Accession: A93780; A01411
R;Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.; Bl
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A;Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing facto
A;Reference number: A93780; MUID:72094314; PMID:4550508
A;Accession: A93780
A;Molecule type: protein
A;Residues: 1-10 <BUR>
A;Note: the natural and synthetic hormones have the same biological activity
C;Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and fo
C;Superfamily: gonadoliberin
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Gly) #status experimental

  Query Match             24.8%;  Score 26;  DB 1;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 9.7e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 SXGCRPG 17
               | | |||
Db          4 SYGLRPG 10


RESULT 5
I78838
flt3 ligand isoform E6 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 05-Nov-1999
C;Accession: I78838
R;Lyman, S.D.; James, L.; Escobar, S.; Downey, H.; de Vries, P.; Brasel, K.; Stocking,
Oncogene 10, 149-157, 1995
A;Title: Identification of soluble and membrane-bound isoforms of the murine flt3 ligan
A;Reference number: I58343; MUID:95124710; PMID:7824267
A;Accession: I78838
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <RES>
A;Cross-references: GB:S76461; NID:g913481; PIDN:AAB33070.1; PID:g913482

  Query Match             23.8%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |     |:||
Db          2 CLEVQCQPG 10


RESULT 6
AKLQIM
locustamyoinhibiting peptide - migratory locust
C;Species: Locusta migratoria (migratory locust)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C;Accession: A60065
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; De Loof, A.
Regul. Pept. 36, 111-119, 1991
A;Title: Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-M
A;Reference number: A60065; MUID:92179466; PMID:1796179
A;Accession: A60065
A;Molecule type: protein
A;Residues: 1-9 <SCH>
C;Comment: This peptide hormone suppresses spontaneous contractions of the hindgut and
C;Superfamily: locustamyoinhibiting peptide
C;Keywords: amidated carboxyl end; hormone
F;9/Modified site: amidated carboxyl end (Trp) #status experimental

  Query Match             22.9%;  Score 24;  DB 1;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
              ||   ||
Db          3 QDLNAGW 9


RESULT 7
D45900
complement C3b receptor type 2 - mouse (clone 12) (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: D45900
R;Kurtz, C.B.; O'Toole, E.; Christensen, S.M.; Weis, J.H.
J. Immunol. 144, 3581-3591, 1990
A;Title: The murine complement receptor gene family. IV. Alternative splicing of Cr2 ge
A;Reference number: A45900; MUID:90229754; PMID:2139460
A;Accession: D45900
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr
A;Molecule type: mRNA
A;Residues: 1-11 <KUR>

  Query Match             22.9%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   37.5%;  Pred. No. 2e+03;
  Matches    3;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRP 16
              |:   | |
Db          2 CEEISCDP 9
```

```
RESULT 8
S39012
proteinase - Thermus sp.
C;Species: Thermus sp.
C;Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 07-May-1999
C;Accession: S39012
R;Freeman, S.A.; Peek, K.; Prescott, M.; Daniel, R.
Biochem. J. 295, 463-469, 1993
A;Title: Characterization of a chelator-resistant proteinase from Thermus strain Rt4A2.
A;Reference number: S39012; MUID:94058984; PMID:8240244
A;Accession: S39012
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRE>
A;Note: 13-Ala was also found

  Query Match             22.9%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 VGWGCD 10
              | || |
Db          6 VTWGLD 11


RESULT 9
A44560
terephthalate 1,2-dioxygenase (EC 1.-.-.-) oxygenase component alpha chain - Comamonas t
C;Species: Comamonas testosteroni
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Feb-1994
C;Accession: A44560
R;Schlaefli, H.
submitted to the Protein Sequence Database, December 1993
A;Description: Terephthalate 1,2-dioxygenase System from Comamonas testosteroni T-2: pur
A;Reference number: A44560
A;Accession: A44560
A;Molecule type: protein
A;Residues: 1-17 <SCH>
A;Note: it is uncertain whether the residue at position 9 is His or Arg
C;Keywords: oxidoreductase

  Query Match             22.9%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   42.9%;  Pred. No. 3e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
              |::|: |
Db          2 QESIIQW 8


RESULT 10
A60421
hypertrehalosemic hormone - German cockroach
N;Alternate names: Bld-HrTH
C;Species: Blattella germanica (German cockroach)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 31-Oct-1997
C;Accession: A60421; S09137
R;Veenstra, J.A.; Camps, F.
Neuropeptides 15, 107-109, 1990
A;Title: Structure of the hypertrehalosemic neuropeptide of the German cockroach, Blatte
A;Reference number: A60421; MUID:91179584; PMID:2080017
A;Accession: A60421
A;Molecule type: protein
A;Residues: 1-10 <VEE>
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659; PMID:2340112
A;Accession: S09137
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

  Query Match             21.9%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 11
S08997
hypertrehalosemic neuropeptide Bld-HrTH - cockroach (Gromphadorina portentosa)
C;Species: Gromphadorina portentosa
C;Date: 30-Jun-1992 #sequence_revision 14-Sep-1994 #text_change 24-Oct-1997
C;Accession: S08997
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpc
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombar
A;Reference number: S08995; MUID:90253659; PMID:2340112
A;Accession: S08997
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

  Query Match             21.9%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 12
S08998
hypertrehalosemic neuropeptide Bld-HrTH - Madeira cockroach
C;Species: Leucophaea maderae (Madeira cockroach)
C;Date: 30-Jun-1992 #sequence_revision 14-Sep-1994 #text_change 24-Oct-1997
C;Accession: S08998
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpc
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombar
A;Reference number: S08995; MUID:90253659; PMID:2340112
A;Accession: S08998
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

  Query Match             21.9%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 13
A26381
hypertrehalosemic hormone - gray cockroach
C;Species: Nauphoeta cinerea (gray cockroach)
```

C;Date: 31-Mar-1988 #sequence_revision 24-Oct-1997 #text_change 31-Oct-1997
C;Accession: A26381
R;Gade, G.; Rinehart Jr., K.L.
Biochem. Biophys. Res. Commun. 141, 774-781, 1986
A;Title: Amino acid sequence of a hypertrehalosaemic neuropeptide from the corpus cardia
A;Reference number: A26381; MUID:87100208; PMID:3801028
A;Accession: A26381
A;Molecule type: protein
A;Residues: 1-10 <GAD>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 21.9%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GWG 8
|||
Db 7 GWG 9

RESULT 14
A31571
hypertrehalosemic/adipokinetic hormone - bollworm
N;Alternate names: Hez-HrTH
C;Species: Heliothis zea (bollworm, corn earworm, tomato fruitworm)
C;Date: 30-Jun-1989 #sequence_revision 23-Mar-1995 #text_change 31-Oct-1997
C;Accession: A31571
R;Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Bird, T.G.; Tseng, C.M.; Zhang, Y.S
Biochem. Biophys. Res. Commun. 155, 344-350, 1988
A;Title: Isolation and primary structure of a neuropeptide hormone from Heliothis zea wi
A;Reference number: A31571; MUID:88326324; PMID:3415690
A;Accession: A31571
A;Molecule type: protein
A;Residues: 1-10 <JAF>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 21.9%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GWG 8
|||
Db 7 GWG 9

RESULT 15
B33995
hypotrehalosemic hormone - black horse fly
C;Species: Tabanus atratus (black horse fly)
C;Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 31-Oct-1997
C;Accession: B33995
R;Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Nachman, R.J.; Vogel, V.W.; Zhang,
Proc. Natl. Acad. Sci. U.S.A. 86, 8161-8164, 1989
A;Title: Primary structure of two neuropeptide hormones with adipokinetic and hypotrehal
A;Reference number: A33995; MUID:90046758; PMID:2813385
A;Accession: B33995
A;Molecule type: protein
A;Residues: 1-10 <JAF>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F;10/Modified site: amidated carboxyl end (Tyr) #status predicted

Query Match 21.9%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GWG 8
|||
Db 7 GWG 9

Search completed: March 13, 2004, 07:26:51
Job time : 22 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:26:14 ; Search time 11 Seconds
(without alignments)
85.206 Million cell updates/sec

Title: US-09-747-029B-17
Perfect score: 105
Sequence: 1 QDTIVGWGCDSXGCRPGQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1002

Minimum DB seq length: 0
Maximum DB seq length: 18

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	30.5	17	1	CXMA_CONPE	P58926 conus penna
2	32	30.5	17	1	CXMB_CONPE	P58927 conus penna
3	25	23.8	17	1	NEF_HV1J3	P12480 human immun
4	24	22.9	9	1	LMIP_LOCMI	P31799 locusta mig
5	24	22.9	15	1	TA1_TREBR	P34070 tremella br
6	23	21.9	10	1	HTF_HELZE	P16353 heliothis z
7	23	21.9	10	1	HTF_NAUCI	P10939 nauphoeta c
8	23	21.9	10	1	HTF_TABAT	P14596 tabanus atr
9	23	21.9	11	1	TKNA_SCYCA	P41333 scyliorhinu
10	23	21.9	17	1	PA2_AUSSU	P59066 austrelaps
11	22	21.0	10	1	GON1_ALLMI	P37041 alligator m
12	22	21.0	11	1	CA31_LITCI	P82089 litoria cit
13	22	21.0	11	1	CA32_LITCI	P82090 litoria cit
14	22	21.0	12	1	HCYB_MEGCR	Q10584 megathura c
15	22	21.0	15	1	NXSO_PSETE	P59073 pseudonaja
16	22	21.0	16	1	LPK1_LOCMI	P20404 locusta mig
17	22	21.0	18	1	SODM_MYCHA	P80582 mycobacteri
18	21	20.0	12	1	CXST_CONTE	P58846 conus texti
19	21	20.0	13	1	CXL4_CONMR	P58810 conus marmo
20	21	20.0	14	1	KPP1_SELMI	P25933 selenastrum
21	21	20.0	15	1	UC19_MAIZE	P80625 zea mays (m
22	21	20.0	16	1	BAII_EUBSP	P32371 eubacterium
23	20	19.0	8	1	ACI_THUAL	P18691 thunnus alb
24	20	19.0	8	1	LCK6_LEUMA	P19988 leucophaea
25	20	19.0	10	1	CA12_LITCI	P82086 litoria cit
26	20	19.0	10	1	CAER_LITXA	P56264 litoria xan
27	20	19.0	10	1	GON1_CLUPA	P81749 clupea pall
28	20	19.0	13	1	BPP1_BOTJA	P01020 bothrops ja
29	20	19.0	14	1	MAST_VESBA	P21654 vespa basal
30	20	19.0	15	1	KPP2_SELMI	P25934 selenastrum
31	20	19.0	16	1	LEC_DELRE	P83511 delonix reg
32	20	19.0	17	1	PATS_ANASP	O52748 anabaena sp
33	19	18.1	12	1	CXA1_CONIM	P50983 conus imper
34	19	18.1	15	1	CX1B_CONBE	P58624 conus betul
35	19	18.1	16	1	TRYP_FELCA	P81071 felis silve
36	19	18.1	17	1	GAST_MACMU	P33714 macaca mula
37	19	18.1	17	1	GPX4_PINPS	P81087 pinus pinas
38	18.5	17.6	15	1	GLN2_PINPS	P81107 pinus pinas
39	18	17.1	8	1	CCKN_MACEU	P30369 macropus eu
40	18	17.1	9	1	FAR5_CALVO	P41860 calliphora
41	18	17.1	10	1	GON2_CHICK	P37043 gallus gall
42	18	17.1	10	1	GON3_ONCKE	P20367 oncorhynchu
43	18	17.1	10	1	GONL_SQUAC	P27429 squalus aca
44	18	17.1	12	1	RF1_CONSP	P58805 conus spuri
45	18	17.1	14	1	IF2G_RAT	P81795 rattus norv

ALIGNMENTS

RESULT 1
CXMA_CONPE
ID CXMA_CONPE STANDARD; PRT; 17 AA.
AC P58926;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mu-conotoxin PnIVA.
OS Conus pennaceus (Feathered cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=37335;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RX MEDLINE=95337083; PubMed=7612605;
RA Fainzilber M., Nakamura T., Gaathon A., Lodder J.C., Kits K.S.,
RA Burlingame A.L., Zlotkin E.;
RT "A new cysteine framework in sodium channel blocking conotoxins.";
RL Biochemistry 34:8649-8656(1995).
CC -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
CC channel. Blocks reversibly sodium channels in molluskan neurons,
CC but has no effect on sodium currents in bovine chromaffin cells or
CC in rat brain synaptosomes. Induces paralysis in bivalve mollusks
CC (Mytilus). No effect are observed on fish (Gambusia) and fly
CC larvae (Sarcophaga). PnIVB is approximately 6 times more potent
CC than PnIVA in blockade of the sodium current in Lymnaea neurons.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1789.5; METHOD=LSIMS.
CC -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
CC FAMILY.
KW Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT SITE 4 4 IMPORTANT FOR BINDING AND ACTIVITY.
FT DISULFID 1 12 BY SIMILARITY.
FT DISULFID 2 15 BY SIMILARITY.
FT DISULFID 8 17 BY SIMILARITY.
SQ SEQUENCE 17 AA; 1797 MW; F9B721E0E96B9D82 CRC64;

Query Match 30.5%; Score 32; DB 1; Length 17;
Best Local Similarity 46.2%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 3; Indels 4; Gaps 1;

```
Qy          6 GW----GCDSXGC 14
              ||    ||  ||
Db          5 GWTCLLGCSPCGC 17
```

RESULT 2
CXMB_CONPE
ID CXMB_CONPE STANDARD; PRT; 17 AA.
AC P58927;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mu-conotoxin PnIVB.

```
OS   Conus pennaceus (Feathered cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=37335;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RX   MEDLINE=95337083; PubMed=7612605;
RA   Fainzilber M., Nakamura T., Gaathon A., Lodder J.C., Kits K.S.,
RA   Burlingame A.L., Zlotkin E.;
RT   "A new cysteine framework in sodium channel blocking conotoxins.";
RL   Biochemistry 34:8649-8656(1995).
RN   [2]
RP   PHARMACOLOGICAL CHARACTERIZATION.
RX   MEDLINE=95346025; PubMed=7620628;
RA   Hasson A., Fainzilber M., Zlotkin E., Spira M.E.;
RT   "Electrophysiological characterization of a novel conotoxin that
RT   blocks molluscan sodium channels.";
RL   Eur. J. Neurosci. 7:815-818(1995).
CC   -!- FUNCTION: Mu-conotoxins bind and block voltage-sensitive sodium
CC       channel. Blocks reversibly sodium channels in molluskan neurons,
CC       but has no effect on sodium currents in bovine chromaffin cells or
CC       in rat brain synaptosomes. Induces paralysis in bivalve mollusks
CC       (Mytilus). No effect are observed on fish (Gambusia) and fly
CC       larvae (Sarcophaga). Is approximately 6 times more potent than
CC       PnIVA in blockade of the sodium current in Lymnaea neurons.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC   -!- MASS SPECTROMETRY: MW=1862.8; METHOD=LSIMS.
CC   -!- SIMILARITY: BELONGS TO THE M-SUPERFAMILY OF CONOTOXINS. MU-TYPE
CC       FAMILY.
KW   Toxin; Neurotoxin; Ionic channel inhibitor; Sodium channel inhibitor.
FT   SITE          4      4       IMPORTANT FOR BINDING AND ACTIVITY (BY
FT                                SIMILARITY).
FT   DISULFID      1     12       BY SIMILARITY.
FT   DISULFID      2     15       BY SIMILARITY.
FT   DISULFID      8     17       BY SIMILARITY.
SQ   SEQUENCE   17 AA;  1870 MW;  E40021E0E96B9D82 CRC64;

  Query Match             30.5%;  Score 32;  DB 1;  Length 17;
  Best Local Similarity   46.2%;  Pred. No. 47;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    4;  Gaps    1;

Qy          6 GW----GCDSXGC 14
              ||    ||   ||
Db          5 GWTCWLGCSPCGC 17


RESULT 3
NEF_HV1J3
ID   NEF_HV1J3      STANDARD;      PRT;    17 AA.
AC   P12480;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Negative factor (F-protein) (27 kDa protein) (3'ORF) (Fragment).
GN   NEF.
OS   Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11694;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89352108; PubMed=2669897;
RA   Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
RT   "Nucleotide sequences of gag and env genes of a Japanese isolate of
RT   HIV-1 and their expression in bacteria.";
RL   AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC   -!- FUNCTION: NEF has GTPase, GTP-binding and autophosphorylating
CC       activities. It seems to down-regulate the CD4(T4) antigen.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21138; AAB03527.1; -.
DR   HIV; M21138; NEF$JH3.
DR   InterPro; IPR001558; HIV_Nef.
DR   Pfam; PF00469; F-protein; 1.
KW   AIDS; Myristate; GTP-binding; Lipoprotein.
FT   LIPID         2      2       N-myristoyl glycine (in host) (By
FT                                similarity).
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1901 MW;  6E6B3F26EFEB921E CRC64;

  Query Match             23.8%;  Score 25;  DB 1;  Length 17;
  Best Local Similarity   60.0%;  Pred. No. 5.8e+02;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 TIVGW 7
              ::|||
Db          9 SVVGW 13


RESULT 4
LMIP_LOCMI
ID   LMIP_LOCMI     STANDARD;      PRT;     9 AA.
AC   P31799;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   Locustamyoinhibiting peptide (LOM-MIP).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92179466; PubMed=1796179;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, identification and synthesis of locustamyoinhibiting
RT   peptide (LOM-MIP), a novel biologically active neuropeptide from
RT   Locusta migratoria.";
RL   Regul. Pept. 36:111-119(1991).
CC   -!- FUNCTION: Suppresses spontaneous contractions of the hindgut and
CC       oviduct.
CC   -!- TISSUE SPECIFICITY: Neurons located in two ventral cell clusters
CC       in the suboesophageal ganglion.
DR   PIR; A60065; AKLQIM.
KW   Amidation; Neuropeptide.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1060 MW;  387D7DD4472AB6C3 CRC64;

  Query Match             22.9%;  Score 24;  DB 1;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+05;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
              ||   ||
Db          3 QDLNAGW 9


RESULT 5
TA1_TREBR
ID   TA1_TREBR      STANDARD;      PRT;    15 AA.
AC   P34070;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tremerogen A-I.
OS   Tremella brasiliensis (Jelly fungus).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
```

```
OC   Tremellomycetidae; Tremellales; Tremellaceae; Tremella.
OX   NCBI_TaxID=29896;
RN   [1]
RP   SEQUENCE.
RA   Ishibashi Y., Sakagami Y., Isogai A., Suzuki A.;
RT   "Structures of Tremerogens A-9291-I and A-9291-VIII: peptidyl sex
RT   hormones of Tremella brasiliensis.";
RL   Biochemistry 23:1399-1404(1984).
CC   -!- FUNCTION: Tremerogen A-I is produced by the a mating-type cells
CC       and induces formation of conjugation tubes in a mating-type cells.
KW   Pheromone; Prenylation; Lipoprotein.
FT   LIPID        15     15       S-farnesyl cysteine.
SQ   SEQUENCE   15 AA;  1339 MW;  3AABA4FC2D605333 CRC64;

  Query Match             22.9%;  Score 24;  DB 1;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          8 GCDSXGC 14
              | | | ||
Db          9 GASSGGC 15


RESULT 6
HTF_HELZE
ID   HTF_HELZE      STANDARD;      PRT;    10 AA.
AC   P16353;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypertrehalosaemic hormone (HeZ-HRTH).
OS   Heliothis zea (Corn earworm) (Bollworm).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC   Noctuidae; Heliothinae; Helicoverpa.
OX   NCBI_TaxID=7113;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=88326324; PubMed=3415690;
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Bird T.G.,
RA   Tseng C.M., Zhang Y.S., Hayes D.K.;
RT   "Isolation and primary structure of a neuropeptide hormone from
RT   Heliothis zea with hypertrehalosemic and adipokinetic activities.";
RL   Biochem. Biophys. Res. Commun. 155:344-350(1988).
CC   -!- FUNCTION: Hypertrehalosaemic factors are neuropeptides that
CC       elevate the level of trehalose in the hemolymph (trehalose is the
CC       major carbohydrate in the hemolymph of insects).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the AKH / HRTH / RPCH family.
DR   PIR; A31571; A31571.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1096 MW;  8E70367865A5B9D1 CRC64;

  Query Match             21.9%;  Score 23;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 7
HTF_NAUCI
ID   HTF_NAUCI      STANDARD;      PRT;    10 AA.
AC   P10939;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypertrehalosaemic hormone (HTH) (Hypertrehalosaemic neuropeptide).
OS   Nauphoeta cinerea (Cinereous cockroach) (Gray cockroach),
OS   Leucophaea maderae (Madeira cockroach),
OS   Blattella germanica (German cockroach), and
OS   Gromphadorina portentosa (Madagascan hissing cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC   Blaberidae; Nauphoeta.
OX   NCBI_TaxID=6990, 6988, 6973, 36953;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=N.cinerea; TISSUE=Corpora cardiaca;
RX   MEDLINE=87100208; PubMed=3801028;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Amino acid sequence of a hypertrehalosaemic neuropeptide from the
RT   corpus cardiacum of the cockroach, Nauphoeta cinerea.";
RL   Biochem. Biophys. Res. Commun. 141:774-781(1986).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=L.maderae, G.portentosa, and B.germanica;
RX   MEDLINE=90253659; PubMed=2340112;
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=B.germanica;
RX   MEDLINE=91179584; PubMed=2080017;
RA   Veenstra J.A., Camps F.;
RT   "Structure of the hypertrehalosemic neuropeptide of the German
RT   cockroach, Blattella germanica.";
RL   Neuropeptides 15:107-109(1990).
CC   -!- FUNCTION: Hypertrehalosaemic factors are neuropeptides that
CC       elevate the level of trehalose in the hemolymph (trehalose is the
CC       major carbohydrate in the hemolymph of insects).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the AKH / HRTH / RPCH family.
DR   PIR; A26381; A26381.
DR   PIR; A60421; A60421.
DR   PIR; S08997; S08997.
DR   PIR; S08998; S08998.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1092 MW;  056236786775B9C4 CRC64;

  Query Match             21.9%;  Score 23;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 8
HTF_TABAT
ID   HTF_TABAT      STANDARD;      PRT;    10 AA.
AC   P14596;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypertrehalosaemic factor (HOTH) (Dipteran corpora cardiaca factor II)
DE   (DCC II).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
```

```
OC   Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha; Tabanidae;
OC   Tabanus.
OX   NCBI_TaxID=7207;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=90046758; PubMed=2813385;
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: Hypertrehalosaemic factors are neuropeptides that
CC       elevate the level of trehalose in the hemolymph (trehalose is the
CC       major carbohydrate in the hemolymph of insects).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the AKH / HRTH / RPCH family.
DR   PIR; B33995; B33995.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1169 MW;  916036786771A9D1 CRC64;

  Query Match             21.9%;  Score 23;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GWG 8
              |||
Db          7 GWG 9


RESULT 9
TKNA_SCYCA
ID   TKNA_SCYCA     STANDARD;      PRT;    11 AA.
AC   P41333;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Substance P.
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=93292508; PubMed=7685693;
RA   Waugh D., Wang Y., Hazon N., Balment R.J., Conlon J.M.;
RT   "Primary structures and biological activities of substance-P-related
RT   peptides from the brain of the dogfish, Scyliorhinus canicula.";
RL   Eur. J. Biochem. 214:469-474(1993).
CC   -!- FUNCTION: Tachykinins are active peptides which excite neurons,
CC       evoke behavioral responses, are potent vasodilators and
CC       secretagogues, and contract (directly or indirectly) many smooth
CC       muscles.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the tachykinin family.
DR   PIR; S33300; S33300.
DR   InterPro; IPR002040; Tachy_Neurokinin.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1278 MW;  214860DEC9D6D867 CRC64;

  Query Match             21.9%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 RPGQ 18
              ||||
Db          3 RPGQ 6


RESULT 10
PA2_AUSSU
ID   PA2_AUSSU      STANDARD;      PRT;    17 AA.
AC   P59066;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospholipase A2 (EC 3.1.1.4) (Phosphatidylcholine 2-acylhydrolase)
DE   (Fragment).
OS   Austrelaps superbus (Australian copperhead).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Austrelaps.
OX   NCBI_TaxID=29156;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Venom;
RX   MEDLINE=93369790; PubMed=8362372;
RA   Yuan Y., Jackson S.P., Mitchell C.A., Salem H.H.;
RT   "Purification and characterisation of a snake venom phospholipase A2:
RT   a potent inhibitor of platelet aggregation.";
RL   Thromb. Res. 70:471-481(1993).
CC   -!- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
CC       acyl groups in 3-sn-phosphoglycerides. Inhibits collagen-, ADP-,
CC       thrombin-, ionophore-, adrenaline-, ristocetin-, and arachidonic
CC       acid-induced platelet aggregation. Inhibits serotonin release.
CC   -!- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
CC       acylglycerophosphocholine + a fatty acid anion.
CC   -!- COFACTOR: Calcium (Probable).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: Belongs to the phospholipase A2 family. Group I
CC       subfamily.
DR   InterPro; IPR001211; PhospholipaseA2.
DR   Pfam; PF00068; phoslip; 1.
KW   Lipid degradation; Hydrolase; Toxin; Calcium.
FT   NON_TER       17     17
SQ   SEQUENCE   17 AA;  1846 MW;  03FB7DD7B7D7D1CB CRC64;

  Query Match             21.9%;  Score 23;  DB 1;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          8 GCDSXGCR 15
              || : | |
Db         10 GCANHGSR 17


RESULT 11
GON1_ALLMI
ID   GON1_ALLMI     STANDARD;      PRT;    10 AA.
AC   P37041; P20407;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gonadoliberin I (Gonadotropin-releasing hormone I) (GnRH-I) (LH-RH I)
DE   (Luliberin I).
OS   Alligator mississippiensis (American alligator).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Crocodylidae; Alligatorinae; Alligator.
OX   NCBI_TaxID=8496;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Brain;
RX   MEDLINE=91352338; PubMed=1882082;
RA   Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
RA   Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT   "Primary structure of two forms of gonadotropin-releasing hormone
```

```
RT   from brains of the American alligator (Alligator mississippiensis).";
RL   Regul. Pept. 33:105-116(1991).
CC   -!- FUNCTION: Stimulates the secretion of gonadotropins.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the GnRH family.
DR   PIR; A60066; RHAQ1.
DR   InterPro; IPR002012; GnRH.
DR   Pfam; PF00446; GnRH; 1.
DR   PROSITE; PS00473; GNRH; 1.
KW   Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1172 MW;  284B23D7286B45A3 CRC64;

  Query Match             21.0%;  Score 22;  DB 1;  Length 10;
  Best Local Similarity   57.1%;  Pred. No. 1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         11 SXGCRPG 17
                | | :||
Db          4 SYGLQPG 10


RESULT 12
CA31_LITCI
ID   CA31_LITCI     STANDARD;      PRT;    11 AA.
AC   P82089;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Caerulein 3.1/3.1Y4.
OS   Litoria citropa (Australian blue mountains tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=94770;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RX   MEDLINE=20057701; PubMed=10589099;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J.;
RT   "Caerulein-like peptides from the skin glands of the Australian blue
RT   montains tree frog Litoria citropa. Part 1. Sequence determination
RT   using electrospray mass spectrometry.";
RL   Rapid Commun. Mass Spectrom. 13:2498-2502(1999).
CC   -!- FUNCTION: Hypotensive neuropeptide (Probable).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin dorsal glands.
CC   -!- PTM: Isoform 3.1Y4 differs from isoform 3.1 in not being
CC       sulfated.
CC   -!- MASS SPECTROMETRY: MW=1407; METHOD=Electrospray.
CC   -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00259; GASTRIN; FALSE_NEG.
KW   Amphibian defense peptide; Hypotensive agent; Amidation; Sulfation;
KW   Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       4      4       SULFATION.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1347 MW;  10DAB7D67861A86B CRC64;

  Query Match             21.0%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
                || |||
Db          2 QDYGTGW 8


RESULT 13
CA32_LITCI
ID   CA32_LITCI     STANDARD;      PRT;    11 AA.
AC   P82090;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Caerulein 3.2/3.2Y4.
OS   Litoria citropa (Australian blue mountains tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=94770;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RX   MEDLINE=20057701; PubMed=10589099;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J.;
RT   "Caerulein-like peptides from the skin glands of the Australian blue
RT   montains tree frog Litoria citropa. Part 1. Sequence determination
RT   using electrospray mass spectrometry.";
RL   Rapid Commun. Mass Spectrom. 13:2498-2502(1999).
CC   -!- FUNCTION: Hypotensive neuropeptide (Probable).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin dorsal glands.
CC   -!- PTM: Isoform 3.2Y4 differs from isoform 3.2 in not being
CC       sulfated.
CC   -!- MASS SPECTROMETRY: MW=1423; METHOD=Electrospray.
CC   -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00259; GASTRIN; FALSE_NEG.
KW   Amphibian defense peptide; Hypotensive agent; Amidation; Sulfation;
KW   Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       4      4       SULFATION.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1363 MW;  10DAB8867861A86B CRC64;

  Query Match             21.0%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
                || |||
Db          2 QDYGTGW 8


RESULT 14
HCYB_MEGCR
ID   HCYB_MEGCR     STANDARD;      PRT;    12 AA.
AC   Q10584;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hemocyanin B chain (KLH-B) (Fragment).
OS   Megathura crenulata (Giant keyhole limpet).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Vetigastropoda; Fissurelloidea; Fissurellidae; Megathura.
OX   NCBI_TaxID=55429;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96208935; PubMed=8829804;
RA   Swerdlow R.D., Ebert R.F., Lee P., Bonaventura C., Miller K.I.;
RT   "Keyhole limpet hemocyanin: structural and functional
RT   characterization of two different subunits and multimers.";
RL   Comp. Biochem. Physiol. 113B:537-548(1996).
CC   -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers
CC       occurring freely dissolved in the hemolymph of many mollusks and
CC       arthropods.
CC   -!- SUBUNIT: Dodecamers and extended multimers.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Hemolymph.
CC   -!- BIOTECHNOLOGY: Potent immunogen used classically as a carrier
CC       protein for haptens and more recently in human vaccines and for
CC       immunotherapy of bladder cancer.
CC   -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin
```

```
CC       subfamily.
DR   InterPro; IPR000896; Hemocyanin.
DR   PROSITE; PS00209; HEMOCYANIN_1; PARTIAL.
DR   PROSITE; PS00210; HEMOCYANIN_2; PARTIAL.
KW   Oxygen transport; Transport; Copper; Glycoprotein;
KW   Hemolymph.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1345 MW;  CBFBEAA44A432412 CRC64;

  Query Match             21.0%;  Score 22;  DB 1;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 DTIVGWGCDS 11
              ||:|     ||
Db          2 DTVVRKNVDS 11


RESULT 15
NXSO_PSETE
ID   NXSO_PSETE     STANDARD;      PRT;    15 AA.
AC   P59073;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Short neurotoxin N2 (Alpha neurotoxin) (Fragment).
OS   Pseudonaja textilis (Eastern brown snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Pseudonaja.
OX   NCBI_TaxID=8673;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=99449602; PubMed=10518793;
RA   Gong N.L., Armugam A., Jeyaseelan K.;
RT   "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: cDNA
RT   cloning, expression and protein characterization.";
RL   Eur. J. Biochem. 265:982-989(1999).
CC   -!- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
CC       acetylcholine receptors (nAChR).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- MASS SPECTROMETRY: MW=6345; METHOD=Electrospray.
CC   -!- MISCELLANEOUS: LD(50) is 0.80 mg/kg by intravenous injection.
CC   -!- SIMILARITY: Belongs to the snake toxin family.
DR   InterPro; IPR003571; Snake_toxin.
DR   PROSITE; PS00272; SNAKE_TOXIN; PARTIAL.
KW   Toxin; Neurotoxin; Postsynaptic neurotoxin;
KW   Acetylcholine receptor inhibitor; Multigene family.
FT   UNSURE       3      3
FT   UNSURE      13     13
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1727 MW;  E149FD4BFD1EF0DD CRC64;

  Query Match             21.0%;  Score 22;  DB 1;  Length 15;
  Best Local Similarity   33.3%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    8;  Gaps    1;

Qy          2 DTIVGWGCDSXGCRP 16
              ||:|       |:|
Db          9 DTVV--------CKP 15


Search completed: March 13, 2004, 07:27:15
Job time : 13 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:26:14 ; Search time 39 Seconds
(without alignments)
145.624 Million cell updates/sec

Title: US-09-747-029B-17
Perfect score: 105
Sequence: 1 QDTIVGWGCDSXGCRPGQ 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 5675

Minimum DB seq length: 0
Maximum DB seq length: 18

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	32.4	15	3	O14379	O14379 schizosacch
2	30	28.6	18	2	Q09053	Q09053 methylomona
3	27	25.7	7	10	O49223	O49223 glycine max
4	27	25.7	18	8	Q8SKY0	Q8sky0 cuscuta ref
5	26	24.8	15	5	Q9TXC8	Q9txc8 locusta mig
6	26	24.8	18	2	Q9R4N5	Q9r4n5 bacillus an
7	25	23.8	17	15	Q9DRS5	Q9drs5 human immun
8	24.5	23.3	16	13	Q9PRY2	Q9pry2 petromyzon
9	24	22.9	9	15	O12096	O12096 caprine art
10	24	22.9	9	15	O12100	O12100 caprine art
11	24	22.9	9	15	O12102	O12102 caprine art
12	24	22.9	9	15	O12098	O12098 caprine art
13	24	22.9	9	15	O12104	O12104 caprine art
14	24	22.9	11	11	Q99JC3	Q99jc3 rattus sp.
15	24	22.9	12	15	O12090	O12090 caprine art
16	24	22.9	12	15	O12094	O12094 caprine art
17	24	22.9	12	15	O12114	O12114 caprine art
18	24	22.9	12	15	O12082	O12082 caprine art
19	24	22.9	12	15	O12106	O12106 caprine art
20	24	22.9	12	15	O12092	O12092 caprine art
21	24	22.9	12	15	O12108	O12108 caprine art
22	24	22.9	12	15	O12074	O12074 caprine art
23	24	22.9	12	15	O12116	O12116 caprine art
24	24	22.9	12	15	O12118	O12118 caprine art
25	24	22.9	12	15	O12110	O12110 caprine art
26	24	22.9	12	15	O12112	O12112 caprine art
27	24	22.9	12	15	O12076	O12076 caprine art
28	24	22.9	12	15	O12088	O12088 caprine art
29	24	22.9	12	15	O12078	O12078 caprine art
30	24	22.9	12	15	O12080	O12080 caprine art
31	24	22.9	12	15	O12084	O12084 caprine art
32	24	22.9	12	15	O12086	O12086 caprine art
33	24	22.9	15	2	Q9R531	Q9r531 thermus. ch
34	24	22.9	17	12	Q85004	Q85004 porcine res
35	23	21.9	9	8	Q94XE6	Q94xe6 tectocoris
36	23	21.9	10	2	Q8KHN9	Q8khn9 clostridium
37	23	21.9	13	8	O99783	O99783 caprimulgus
38	23	21.9	14	4	Q13022	Q13022 homo sapien
39	23	21.9	15	4	Q9UC67	Q9uc67 homo sapien
40	23	21.9	15	8	Q8SJ19	Q8sj19 phalacrocor
41	23	21.9	16	4	Q9UC54	Q9uc54 homo sapien
42	23	21.9	16	6	Q95MB4	Q95mb4 equus cabal
43	23	21.9	16	11	Q80WI5	Q80wi5 mus sp. ago
44	23	21.9	17	8	Q85UD6	Q85ud6 conger myri
45	23	21.9	17	11	P97758	P97758 mus musculu

ALIGNMENTS

RESULT 1
O14379
ID O14379 PRELIMINARY; PRT; 15 AA.
AC O14379;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RA Jang Y.-J., Yoo H.-S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U97375; AAB63867.1; -.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1636 MW; 93D127B36BEAF110 CRC64;

Query Match 32.4%; Score 34; DB 3; Length 15;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          5 VGWGCDSXGCR 15
              | :|| |  ||
Db          4 VDYGCMSLSCR 14
```

RESULT 2
Q09053
ID Q09053 PRELIMINARY; PRT; 18 AA.
AC Q09053;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```
DE   Methanol dehydrogenase subunit 2 (EC 1.1.99.8) (MDH small beta
DE   subunit) (Fragment).
GN   MOXI.
OS   Methylomonas SP.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC   Methylococcaceae; Methylomonas.
OX   NCBI_TaxID=418;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=A4;
RX   MEDLINE=93285990; PubMed=7685335;
RA   Waechter-Brulla D., DiSpirito A.A., Chistoserdova L.V., Lidstrom M.E.;
RL   J. Bacteriol. 175:3767-3775(1993).
CC   -!- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
CC       REDUCED ACCEPTOR.
CC   -!- COFACTOR: PQQ.
CC   -!- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION OF
CC       A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
DR   GO; GO:0018468; F:alcohol dehydrogenase (acceptor) activity; IEA.
DR   GO; GO:0004022; F:alcohol dehydrogenase activity; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0015946; P:methanol oxidation; IEA.
DR   InterPro; IPR003420; Meth_dh_beta.
DR   Pfam; PF02315; MDH; 1.
KW   Oxidoreductase; PQQ; Methanol utilization; Periplasmic.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2014 MW;  2B0412DBD2B8B52D CRC64;

  Query Match             28.6%;  Score 30;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy         10 DSXGCRPG 17
              |   |:||
Db          2 DGTNCKPG 9


RESULT 3
O49223
ID   O49223      PRELIMINARY;      PRT;     7 AA.
AC   O49223;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HMG-1-like protein (Fragment).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Essex; TISSUE=Root;
RX   MEDLINE=91367679; PubMed=1891369;
RA   Laux T., Goldberg R.B.;
RT   "A plant DNA binding protein shares highly conserved sequence motifs
RT   with HMG-box proteins.";
RL   Nucleic Acids Res. 19:4769-4769(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Essex; TISSUE=Root;
RA   Mahalingam R., Knap H.T.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF047050; AAC03556.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  850 MW;  6AAAAAB378637810 CRC64;

  Query Match             25.7%;  Score 27;  DB 10;  Length 7;
  Best Local Similarity   80.0%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 GWGCD 10
              ||| |
Db          1 GWGWD 5


RESULT 4
Q8SKY0
ID   Q8SKY0      PRELIMINARY;      PRT;    18 AA.
AC   Q8SKY0;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein S11 (Fragment).
GN   RPS11.
OS   Cuscuta reflexa (Southern Asian dodder).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Convolvulaceae; Cuscuta.
OX   NCBI_TaxID=4129;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Berg S.;
RT   "Sequence analysis and coding potential of the holoparasitic flowering
RT   plant genus Cuscuta.";
RL   Thesis (2002), Department of Institute of Botany,.
DR   EMBL; AJ439611; CAD28796.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0005622; C:intracellular; IEA.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR001971; Ribosomal_S11.
DR   Pfam; PF00411; Ribosomal_S11; 1.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   18 AA;  2088 MW;  130D427BFE680B24 CRC64;

  Query Match             25.7%;  Score 27;  DB 8;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 GCRP 16
              ||||
Db          9 GCRP 12


RESULT 5
Q9TXC8
ID   Q9TXC8      PRELIMINARY;      PRT;    15 AA.
AC   Q9TXC8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   PROPHENOLOXIDASE inhibitor N terminus (Fragment).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91379003; PubMed=1910340;
RA   Brehelin M., Boigegrain R.A., Drif L., Coletti-Previero M.A.;
RT   "Purification of a protease inhibitor which controls prophenoloxidase
RT   activation in hemolymph of Locusta migratoria (insecta).";
RL   Biochem. Biophys. Res. Commun. 179:841-846(1991).
DR   InterPro; IPR009041; PMP_inhibitor.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1707 MW;  C690EAE0112166C7 CRC64;

  Query Match             24.8%;  Score 26;  DB 5;  Length 15;
```

```
  Best Local Similarity   80.0%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         14 CRPGQ 18
              | |||
Db          4 CTPGQ 8


RESULT 6
Q9R4N5
ID   Q9R4N5      PRELIMINARY;      PRT;    18 AA.
AC   Q9R4N5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   EXTRACTABLE antigen 1 (Fragment).
OS   Bacillus anthracis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1392;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95247684; PubMed=7730281;
RA   Farchaus J.W., Ribot W.J., Downs M.B., Ezzell J.W.;
RT   "Purification and characterization of the major surface array protein
RT   from the avirulent Bacillus anthracis Delta Sterne-1.";
RL   J. Bacteriol. 177:2481-2489(1995).
SQ   SEQUENCE   18 AA;  1926 MW;  1DBEBF0A4925EFB6 CRC64;

  Query Match             24.8%;  Score 26;  DB 2;  Length 18;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 WGCDS 11
              || ||
Db         13 WGIDS 17


RESULT 7
Q9DRS5
ID   Q9DRS5      PRELIMINARY;      PRT;    17 AA.
AC   Q9DRS5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Truncated nef protein (Negative factor) (27 kDa protein).
GN   NEF.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LTS 25e;
RA   Ashton L., Rhodes D., Solomon A., Deacon N., Satchell C., Carr A.,
RA   Cooper D., Biti R., Stewart G., Kaldor J.;
RT   "Viral diversity in the nef/LTR region of the HIV-1 genome:
RT   associations with long-term nonprogression.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: NEF HAS GTPASE, GTP-BINDING AND AUTOPHOSPHORYLATING
CC       ACTIVITIES. IT SEEMS TO DOWN-REGULATE THE CD4(T4) ANTIGEN (BY
CC       SIMILARITY).
DR   EMBL; AF219708; AAG44185.1; -.
DR   GO; GO:0005525; F:GTP binding; IEA.
DR   InterPro; IPR001558; HIV_Nef.
DR   Pfam; PF00469; F-protein; 1.
KW   AIDS; GTP-binding; Lipoprotein; Myristate.
SQ   SEQUENCE   17 AA;  1846 MW;  6E6A7A26EFEB808E CRC64;

  Query Match             23.8%;  Score 25;  DB 15;  Length 17;
  Best Local Similarity   28.6%;  Pred. No. 3.8e+03;
  Matches    2;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QDTIVGW 7
              : :::||
Db          7 KSSVIGW 13


RESULT 8
Q9PRY2
ID   Q9PRY2      PRELIMINARY;      PRT;    16 AA.
AC   Q9PRY2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Lectin 30 kDa subunit (Fragment).
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94249896; PubMed=8192354;
RA   Schluter S.F., Schroeder J., Wang E., Marchalonis J.J.;
RT   "Recognition molecules and immunoglobulin domains in invertebrates.";
RL   Ann. N. Y. Acad. Sci. 712:74-81(1994).
SQ   SEQUENCE   16 AA;  1728 MW;  3BBF03DD4185F446 CRC64;

  Query Match             23.3%;  Score 24.5;  DB 13;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          7 WGCDSXGC 14
              | | : ||
Db          1 WSC-TKGC 7


RESULT 9
O12096
ID   O12096      PRELIMINARY;      PRT;     9 AA.
AC   O12096;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81439; AAB60832.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;

  Query Match             22.9%;  Score 24;  DB 15;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9


RESULT 10
O12100
ID   O12100      PRELIMINARY;      PRT;     9 AA.
AC   O12100;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
```

```
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81441; AAB60836.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;

  Query Match             22.9%;  Score 24;  DB 15;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9


RESULT 11
O12102
ID   O12102      PRELIMINARY;      PRT;     9 AA.
AC   O12102;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81442; AAB60838.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;

  Query Match             22.9%;  Score 24;  DB 15;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9


RESULT 12
O12098
ID   O12098      PRELIMINARY;      PRT;     9 AA.
AC   O12098;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81440; AAB60835.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;

  Query Match             22.9%;  Score 24;  DB 15;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9


RESULT 13
O12104
ID   O12104      PRELIMINARY;      PRT;     9 AA.
AC   O12104;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81443; AAB60840.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  922 MW;  21E8644EB7340EB8 CRC64;

  Query Match             22.9%;  Score 24;  DB 15;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1e+06;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9


RESULT 14
Q99JC3
ID   Q99JC3      PRELIMINARY;      PRT;    11 AA.
AC   Q99JC3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Luteinizing hormone/chorionic gonadotropin receptor homolog
DE   (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=96147985; PubMed=8571710;
RA   Shen Q.X., Liu H.H., Chen W.Y., Bahl O.P.;
RT   "[Cloning and overexpression of rat ovary LH/hCG receptor cDNA in
RT   insect cells].";
RL   Shih Yen Sheng Wu Hsueh Pao 28:283-290(1995).
DR   EMBL; S80658; AAB50709.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005213; F:structural constituent of chorion (sensu In. . .; IEA.
KW   Chorion; Receptor.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  994 MW;  333DCB137EB865B8 CRC64;

  Query Match             22.9%;  Score 24;  DB 11;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 3.5e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          6 GWGCDSXG 13
              | ||  : |
Db          4 GSGCGAAG 11
```

RESULT 15
O12090
ID   O12090   PRELIMINARY;   PRT;   12 AA.
AC   O12090;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Tat protein (Fragment).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11660;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.-F., Vigne R., Querat G.;
RT   "dUTPase minus CAEV is attenuated for pathogenesis and accumulates G
RT   to A substitutions.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81436; AAB60826.1; -.
DR   GO; GO:0016563; F:transcriptional activator activity; IEA.
DR   GO; GO:0045941; P:positive regulation of transcription; IEA.
DR   InterPro; IPR004247; Lentiviral_Tat.
DR   Pfam; PF02998; Lentiviral_Tat; 1.
FT   NON_TER   1   1
SQ   SEQUENCE   12 AA; 1266 MW; 5A60BBB1E8644EB7 CRC64;

```
  Query Match             22.9%;  Score 24;  DB 15;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 3.8e+03;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          9 CDSXGCRPG 17
              |    | ||
Db          1 CGCRLCNPG 9
```

Search completed: March 13, 2004, 07:29:49
Job time : 42 secs